UNITED STATES DISTRICT COURT
DISTRICT OF MASSACHUSETTS

UNITED STATES OF AMERICA **v.** **AARON SWARTZ,** **Defendant**	**Crim. No. 11-CR-10260-NMG** **VIOLATIONS:** **18 U.S.C. § 1343 (Wire Fraud)** **18 U.S.C. § 1030(a)(4),(b) (Computer Fraud)** **18 U.S.C. § 1030(a)(2), (b), (c)(2)(B)(iii) (Unlawfully Obtaining Information from a Protected Computer)** **18 U.S.C. § 1030(a)(5)(B), (c)(4)(A)(i)(I),(VI) (Recklessly Damaging a Protected Computer)** **18 U.S.C. § 2 (Aiding and Abetting)** **18 U.S.C. § 981(a)(1)(C), 28 U.S.C. § 2461(c), 18 U.S.C. § 982(a)(2)(B), and 18 U.S.C. § 1030(i) (Criminal Forfeiture)**

SUPERSEDING INDICTMENT

The Grand Jury charges that at all relevant times:

PARTIES

JSTOR

1. JSTOR, founded in 1995, was and continued to be a United States-based, not-for-profit organization that provides an online system for archiving and providing access to academic journals and journal articles. It provides searchable digitized copies of articles from over 1,000 academic journals, dating back for lengthy periods of time.

2. JSTOR's service is important to research institutions and universities because it can be extraordinarily expensive, in terms of both cost and space, for a research or university library to maintain a comprehensive collection of academic journals. By digitizing extensive, historical collections of journals, JSTOR enables libraries to outsource the journals' storage, ensures their preservation, and enables authorized users to conduct full-text, cross-disciplinary

searches of them. JSTOR has invested millions of dollars in obtaining and digitizing the journal articles that it makes available as part of its service.

3. JSTOR generally charges libraries, universities, and publishers a subscription fee for access to JSTOR's digitized journals. For a large research university, this annual subscription fee for JSTOR's various collections of content can cost more than $50,000. Portions of the subscription fees are shared with the journal publishers who hold the original copyrights. In addition, JSTOR makes some articles available for individual purchase.

4. JSTOR authorizes users to download a limited number of journal articles at a time. Before being given access to JSTOR's digital archive, each user must agree and acknowledge that they cannot download or export content from JSTOR'S computer servers with automated computer programs such as web robots, spiders, and scrapers. JSTOR also uses computerized measures to prevent users from downloading an unauthorized number of articles using automated techniques.

MIT

5. The Massachusetts Institute of Technology ("MIT") was and continued to be a leading research and teaching university located in Cambridge, Massachusetts.

6. JSTOR provided MIT with its services and content for a fee.

7. MIT made JSTOR's services and content available to its students, faculty, and employees. MIT also allowed guests of the Institute to have the same access to JSTOR, but required guests to register on the MIT network. MIT authorized guests to use its network for no more than fourteen days per year, and required all users to use the network to support MIT's research, education, and administrative activities, or at least to not interfere with these activities; to maintain the system's security and conform to applicable laws, including copyright laws; and to conform with rules imposed by any networks to which users connected through MIT's system. These rules explicitly notified users that violations could lead to state or federal prosecution. Guest users of the MIT network agreed to be bound by the same rules that applied to students,

faculty, and employees.

8. JSTOR's computers were located outside the Commonwealth of Massachusetts, and thus any communications between JSTOR's computers and MIT's computers crossed state boundaries. JSTOR's and MIT's computers were also used in and affected interstate and foreign commerce.

Aaron Swartz

9. Aaron Swartz lived in the District of Massachusetts and was a fellow at Harvard University's Safra Center for Ethics. Swartz was not affiliated with MIT as a student, faculty member, or employee or in any other manner. Although Harvard provided Swartz access to JSTOR's services and archive as needed for his research, Swartz used MIT's computer networks to steal millions of articles from JSTOR.

OVERVIEW OF THE OFFENSES

10. Between September 24, 2010, and January 6, 2011, Swartz contrived to:

a. break into a restricted-access computer wiring closet at MIT;

b. access MIT's network without authorization from a switch within that closet;

c. access JSTOR's archive of digitized journal articles through MIT's computer network;

d. use this access to download a substantial portion of JSTOR's total archive onto his computers and computer hard drives;

e. avoid MIT's and JSTOR's efforts to prevent this massive copying, efforts that were directed at users generally and at Swartz's illicit conduct specifically; and

f. elude detection and identification.

MEANS OF COMMITTING THE OFFENSES

11. Swartz alone, or in knowing concert with others unknown to the Grand Jury, (hereafter simply "Swartz" in this section) committed these offenses through the means described below.

September 24 through 27, 2010

12. On September 24, 2010, Swartz purchased an Acer laptop computer from a local computer store.

13. Later that day, Swartz connected the Acer laptop to MIT's computer network from a location in Building 16 at MIT and registered with MIT's computer network as a guest.

14. When Swartz registered on the network, he took measures to hide his identity as the computer's owner and user:

a. Swartz registered the computer under the fictitious guest name "Gary Host."

b. Swartz specified the computer's client name as "ghost laptop." (A computer's client name helps to identify it on a network and can be chosen by its user.) In this case, the "ghost" client name abridged the pseudonym "Gary Host" by combining the first initial "g" with the last name "host."

c. Swartz identified the fictitious "Gary Host's" e-mail address as "ghost@mailinator.com", a temporary e-mail address. Mailinator advertised itself as a free e-mail service that allows a user to create a new temporary-mail address as needed. Mailinator advertised that it would accept mail for any e-mail address directed to the mailinator.com domain without need for a prior registration or account. Mailinator also advertised that all mail sent to mailinator.com would automatically be deleted after several hours, whether read or not, and that the company kept no logs of e-mail access.

15. On September 25, 2010, Swartz used the Acer laptop to systematically access and

rapidly download an extraordinary volume of articles from JSTOR by submitting download requests faster than a human could type, and in a manner designed to sidestep or confuse JSTOR's computerized efforts to restrict the volume of individual users' downloads.

16. The effect of these rapid and massive downloads and download requests was to impair computers used by JSTOR to provide articles to client research institutions.

17. As JSTOR, and then MIT, became aware of these events, each took steps to block communications to and from Swartz's computer. Swartz, in turn, altered the apparent source of his automated demands to sidestep or circumvent JSTOR's and MIT's blocks against his computer, as described below:

a. On the evening of September 25, 2010, JSTOR terminated Swartz's computer's network access by refusing communications from the computer's assigned IP address.

i. An IP (short for "Internet Protocol") address is a unique numeric address assigned to each computer connected to the Internet so that the computer's incoming and outgoing Internet traffic is directed to the proper destination. Most Internet service providers control a range of IP addresses. MIT controls all IP addresses that begin with the number 18.

ii. Swartz's computer had been assigned an IP address of 18.55.6.215.

iii. On September 25, 2010, JSTOR blocked communications from that IP address, thus preventing Swartz from requesting and receiving any more JSTOR articles.

b. On September 26, 2010, Swartz established a new IP address for his computer on the MIT network – 18.55.6.216 – which sidestepped the IP address block and allowed the laptop to resume downloading an extraordinary volume of articles from JSTOR. Accesses from this address continued until the middle of the day, when JSTOR spotted the access and blocked communications from this

new IP address as well.

c. Because the downloads on September 25 and 26 originated from shifting MIT IP addresses beginning with 18.55.6, and because JSTOR's computers used to provide articles to research institutions had been impaired and significant portions of its archive was at risk of misappropriation, on September 26, 2010, JSTOR began blocking a broader range of IP addresses. The block prevented a researcher assigned any one of over 250 other IP addresses available at MIT from being able to access JSTOR's archive until September 29, 2010.

d. After JSTOR notified MIT what was happening, MIT sought to block Swartz in particular. It did so by prohibiting Swartz's laptop from being assigned any IP address on MIT's network. MIT did so by blocking communications with any computer bearing the laptop's MAC address.

i. A MAC address is a unique identifier assigned to each computer's network interface, in this case, Swartz's Acer laptop's network interface card.

ii. When a user plugs his computer into MIT's wired network on campus, the network reads the computer's MAC address to determine whether the computer is authorized to use the network. As part of the registration process, "Gary Host's" computer, i.e., Swartz's Acer laptop, had identified its network interface's MAC address as 00:23:5a:73:5f:fb.

iii. Consequently, on September 27, 2010, MIT terminated the laptop's guest registration and barred any network interface with that MAC address from obtaining a new IP address.

October 2 through 9, 2010

18. On October 2, 2010, just over a week after JSTOR and MIT had blocked Swartz's Acer laptop from communicating with JSTOR's and MIT's networks, Swartz sought and

obtained another guest connection on MIT's network for his Acer laptop.

19. Once again, Swartz registered the Acer laptop on the network using identifiers chosen to avoid identifying Swartz as the computer's owner and user:

a. Swartz once again registered the computer under the fictitious name "Gary Host" and the client name "ghost laptop."

b. To evade the MAC address block, Swartz "spoofed" the Acer laptop's computer's MAC address. A MAC address is usually assigned to a network interface card by the card's manufacturer, and therefore generally remains constant. But a user with the right knowledge can change the MAC address, an action referred to as "MAC address spoofing." Swartz spoofed the Acer laptop's MAC address by changing it from 00:23:5a:73:5f:fb to 00:23:5a:73:5f:fc (that is, the final 'b' became a 'c').

c. By re-registering the laptop, the laptop received a new IP address, which disassociated Swartz's Acer laptop from the IP addresses that JSTOR had blocked when Swartz had used them in September.

20. On October 8, 2010, Swartz connected a second computer to MIT's network and registered as a guest, using similar naming conventions: Swartz registered the computer under the name "Grace Host," the computer client name "ghost macbook," and the temporary e-mail address "ghost42@mailinator.com."

21. On October 9, 2010, Swartz used both the "ghost laptop" and the "ghost macbook" to, again, systematically and rapidly access and download articles from JSTOR. The pace of Swartz's automated downloads was so fast and voluminous that it significantly impaired the operation of some computers at JSTOR.

22. In response, beginning on or about October 9, 2010, JSTOR blocked MIT's entire computer network from accessing JSTOR. The block lasted several days, again depriving legitimate users at MIT from accessing JSTOR's services.

November and December, 2010

23. During November and December, 2010, Swartz again used the "ghost laptop" (i.e., the Acer laptop) at MIT to download over two million documents from JSTOR, more than one hundred times the number of downloads during the same period by all legitimate MIT JSTOR users combined.

24. During this period, when Swartz connected to MIT's computer network, he circumvented MIT's guest registration process altogether. Rather than let MIT assign his computer an IP address automatically, Swartz instead simply hard-wired into the network and assigned himself two IP addresses. He did so by entering a restricted network interface closet in the basement of MIT's Building 16, plugging the computer directly into the network, and operating the computer to assign itself two IP addresses. To further cloak his activities, Swartz also hid the Acer laptop and a succession of external storage drives under a box in the closet, so that they would not arouse the suspicions of anyone who might enter the closet.

January 4 through 6, 2011

25. On January 4, 2011, Swartz entered the restricted basement network wiring closet and replaced an external hard drive attached to the laptop.

26. On January 6, 2011, Swartz returned to the wiring closet to remove his computer equipment. This time he attempted to evade identification at the entrance to the restricted area. Apparently aware of or suspicious of a video camera, as Swartz entered the wiring closet, he held his bicycle helmet like a mask to shield his face, looking through ventilation holes in the helmet. Swartz then removed his computer equipment from the closet, put it in his backpack, and left, again masking his face with the bicycle helmet before peering through a crack in the double doors and cautiously stepping out.

27. Later that day, Swartz connected his Acer laptop to MIT's network in a different building — the student center — again registering on the network using identifiers chosen to avoid identifying Swartz as the computer's owner and user:

a. Swartz registered the computer under the fictitious name "Grace Host" and the client name "ghost laptop."

b. By re-registering the laptop, the laptop again received a new IP address, which disassociated Swartz's Acer laptop from the IP addresses Swartz had used up to that point.

c. To again evade the MAC address block, Swartz had spoofed the Acer laptop's MAC address a second time, changing it from the blocked 00:23:5a:73:5f:fb (or from the later-spoofed 00:23:5a:73:5f:fc) to 00:4c:e5:a0:c7:56.

28. Swartz's Acer laptop contained a software program named "keepgrabbing.py," which was designed to download .pdf files (the format used by JSTOR) from JSTOR and sidestep or confuse JSTOR's computerized efforts to prevent repeated and voluminous downloads.

29. When MIT Police spotted Swartz on the afternoon of January 6, 2011 and attempted to question him, Swartz fled with a USB drive that contained the program "keepgrabbing2.py," which was similar to "keepgrabbing.py."

30. In all, Swartz stole a major portion of the total archive in which JSTOR had invested.

31. Swartz intended to distribute these articles through one or more file-sharing sites.

COUNTS 1 and 2
Wire Fraud
18 U.S.C. §§ 1343 & 2

32. The Grand Jury realleges and incorporates by reference the allegations in paragraphs 1-31 of this Indictment.

33. Aaron Swartz devised a scheme to defraud JSTOR of a substantial number of journal articles which they had invested in collecting, obtaining the rights to distribute, and digitizing.

34. He sought to defraud MIT and JSTOR of rights and property by:

a. Deceptively making it appear to JSTOR that he was affiliated with MIT by downloading JSTOR's articles through MIT's computer network and from MIT IP addresses, even though he was not affiliated at the time with MIT, and even though for legitimate research he could have accessed JSTOR through Harvard University, where he worked;

b. Repeatedly taking steps to change his and his computer's apparent identities and to conceal his and his computer's true identities;

c. Using a rapid, automated collection software tool designed to make it appear as if he were multiple people making single requests rather than a single person making multiple requests, in order to bypass safeguards designed to limit the number of articles any one person could download;

d. Attempting to conceal from MIT the physical location of the Acer laptop's connection to MIT's network, by placing it in a utility closet, covering it with cardboard, and, at one point, moving it from one MIT building to another; and

e. Using wire communications between a MIT computer in Massachusetts and JSTOR's computer out-of-state to effectuate his scheme.

35. The Grand Jury charges that repeatedly from on or about September 24, 2010 through January 6, 2011, or thereabout, in the District of Massachusetts and elsewhere, the

defendant,

AARON SWARTZ,

having devised and intended to devise a scheme and artifice to defraud and for obtaining property — journal articles digitized and distributed by JSTOR, and copies of them — by means of material false and fraudulent pretenses and representations, transmitted and caused to be transmitted by means of wire communication in interstate commerce writings, signs, and signals — that is, communications to and from JSTOR's computer servers —for the purpose of executing the scheme, and aiding and abetting it, including on or about the dates specified below:

COUNT	DATES
1	October 9, 2010
2	January 4-6, 2011

All in violation of Title 18, United States Code, Sections 1343 and 2.

COUNTS 3-7
Computer Fraud
18 U.S.C. §§ 1030(a)(4), (b) & 2

36. The Grand Jury realleges and incorporates by reference the allegations in paragraphs 1-31 and 33-34 of this Indictment and charges that:

37. Repeatedly between on or about September 26, 2010 and January 6, 2011, including on or about the dates specified below, in the District of Massachusetts and elsewhere, the defendant,

AARON SWARTZ,

knowingly and with intent to defraud, accessed protected computers belonging to MIT and JSTOR without authorization, and by means of such conduct furthered the intended fraud and obtained things of value — namely, digitized journal articles from JSTOR's archive — and aided and abetted the same and attempted to do the same:

COUNT	DATES	PROTECTED COMPUTERS
3	September 26, 2010	JSTOR
4	October 2-9, 2010	MIT
5	November 29, 2010 - December 26, 2010	JSTOR
6	December 27, 2010 - January 4, 2011	JSTOR
7	January 4-6, 2011	MIT

All in violation of Title 18, United States Code, Sections 1030(a)(4) and 2.

COUNT 8-12
Unlawfully Obtaining Information from a Protected Computer
18 U.S.C. §§ 1030(a)(2), (b), (c)(2)(B)(iii) & 2

38. The Grand Jury realleges and incorporates by reference the allegations in paragraphs 1-31 and 33-34 of this Indictment and charges that:

39. Repeatedly between September 26, 2010 and January 6, 2011, including on or about the dates specified below, in the District of Massachusetts and elsewhere, the defendant,

AARON SWARTZ,

intentionally accessed computers belonging to MIT and JSTOR without authorization, and thereby obtained from protected computers information whose value exceeded $5,000 — namely, digitized journal articles from JSTOR's archive — and aided and abetted the same and attempted to do the same.

COUNT	DATES	PROTECTED COMPUTERS
8	September 26, 2010	JSTOR
9	October 2-9, 2010	MIT
10	November 29, 2010 - December 26, 2010	JSTOR
11	December 27, 2010 - January 4, 2011	JSTOR
12	January 4-6, 2011	MIT

All in violation of 18 U.S.C. §§ 1030(a)(2), (c)(2)(B)(iii) and 2.

COUNT 13
Recklessly Damaging a Protected Computer
18 U.S.C. §§ 1030(a)(5)(B), (c)(4)(A)(i)(I),(VI) & 2

40. The Grand Jury realleges and incorporates by reference the allegations in paragraphs 1-31 and 33-34 of this Indictment.

41. Aaron Swartz's repeated accessing of JSTOR's and MIT's computer systems without authorization constituted a related course of conduct lasting from on or about September 26, 2010 through January 6, 2011. His unauthorized access of the systems on or about days such as September 26 and October 9, 2010 resulted in reckless damage to both. The pace and volume of his automated requests impaired computers JSTOR used to provide service to researchers and research institutions and caused JSTOR to cut off legitimate MIT researchers for days at a time.

42. Both MIT and JSTOR were required to expend significant resources to respond to Swartz's unlawful access to their systems and high speed automated downloads of substantial portions of JSTOR's digital archives.

43. The Grand Jury charges that on or about October 9, 2010 in the District of Massachusetts and elsewhere, the defendant,

AARON SWARTZ,

intentionally accessed a protected computer without authorization, and as a result of such conduct recklessly caused damage to MIT and JSTOR, that is impairment to the availability of information, data, and a system, which, during a one year period:

(A) caused loss, that is, reasonable costs of responding to the offense, conducting a damage assessment, and restoring the information, data, and system to its condition prior to the offense, aggregating at least $5,000 in value from a related course of conduct affecting at least one other protected computer, and

(B) damage affecting at least 10 protected computers.

All in violation of Title 18, United States Code, Section 1030(a)(5)(B), (c)(4)(A)(i)(I),(VI) & 2.

FORFEITURE ALLEGATIONS
(18 U.S.C. § 981(a)(1)(C), 28 U.S.C. § 2461(c), 18 U.S.C. § 982(a)(2)(B), and 18 U.S.C. § 1030(i))

44. Upon conviction of one or more of the offenses alleged in Counts One and Two of the Indictment, the defendant,

AARON SWARTZ,

shall forfeit to the United States, pursuant to 18 U.S.C. § 981(a)(1)(C) and 28 U.S.C. § 2461(c), any property, real or personal, that constitutes, or is derived from, proceeds traceable to the commission of the offense.

45. Upon conviction of one or more of the offenses alleged in Counts Three through Thirteen of the Indictment, the defendant,

AARON SWARTZ,

shall forfeit to the United States, pursuant to 18 U.S.C. § 982(a)(2)(B) and 18 U.S.C. § 1030(i) any property constituting, or derived from, proceeds obtained directly or indirectly as a result of the commission of the offenses, and pursuant to 18 U.S.C. § 1030(i) any personal property that was used or intended to be used to commit or facilitate the commission of such violations.

46. If any of the property described in paragraphs 44 and 45 hereof as being forfeitable pursuant to 18 U.S.C. § 981(a)(1)(C), 28 U.S.C. § 2461(c), 18 U.S.C. § 982(a)(2)(B), and 18 U.S.C. § 1030(i) as a result of any act or omission of the defendant --

a. cannot be located upon the exercise of due diligence;

b. has been transferred to, sold to, or deposited with a third party;

c. has been placed beyond the jurisdiction of this Court;

d. has been substantially diminished in value; or

e. has been commingled with other property which cannot be divided without difficulty;

it is the intention of the United States, pursuant to 21 U.S.C. § 853(p), as incorporated by 28 U.S.C. § 2461(c), 18 U.S.C. § 982(b)(1), and 18 U.S.C. § 1030(i)(2), to seek forfeiture of all

other property of the defendant up to the value of the property described in paragraphs 44 and 45 above.

All pursuant to Title 18, United States Code, Sections 981(a)(1)(C), 982(a)(2)(B), and 1030(i), and Title 28, United States Code, Section 2461(c).

A TRUE BILL

Brenda L Hannon

Foreperson of the Grand Jury

Scott L Garland

Assistant United States Attorney

Date: 9-12-12

DISTRICT OF MASSACHUSETTS

September 12, 2012

Returned into the District Court by the Grand Jurors and filed.

Steve York

Deputy Clerk

12:45 9/12/12